4. The Dirac Equation

> "A great deal more was hidden in the Dirac equation than the author had expected when he wrote it down in 1928. Dirac himself remarked in one of his talks that his equation was more intelligent than its author. It should be added, however, that it was Dirac who found most of the additional insights."
>
> *Weisskopf on Dirac*

So far we've only discussed scalar fields such that under a Lorentz transformation $x^\mu \rightarrow (x')^\mu = \Lambda^\mu_\nu x^\nu$, the field transforms as

$$\phi(x) \rightarrow \phi'(x) = \phi(\Lambda^{-1}x) \tag{4.1}$$

We have seen that quantization of such fields gives rise to spin 0 particles. But most particles in Nature have an intrinsic angular momentum, or spin. These arise naturally in field theory by considering fields which themselves transform non-trivially under the Lorentz group. In this section we will describe the Dirac equation, whose quantization gives rise to fermionic spin 1/2 particles. To motivate the Dirac equation, we will start by studying the appropriate representation of the Lorentz group.

A familiar example of a field which transforms non-trivially under the Lorentz group is the vector field $A_\mu(x)$ of electromagnetism,

$$A^\mu(x) \rightarrow \Lambda^\mu{}_\nu A^\nu(\Lambda^{-1}x) \tag{4.2}$$

We'll deal with this in Section 6. (It comes with its own problems!). In general, a field can transform as

$$\phi^a(x) \rightarrow D[\Lambda]^a{}_b\,\phi^b(\Lambda^{-1}x) \tag{4.3}$$

where the matrices $D[\Lambda]$ form a *representation* of the Lorentz group, meaning that

$$D[\Lambda_1]D[\Lambda_2] = D[\Lambda_1\Lambda_2] \tag{4.4}$$

and $D[\Lambda^{-1}] = D[\Lambda]^{-1}$ and $D[1] = 1$. How do we find the different representations? Typically, we look at infinitesimal transformations of the Lorentz group and study the resulting Lie algebra. If we write,

$$\Lambda^\mu{}_\nu = \delta^\mu{}_\nu + \omega^\mu{}_\nu \tag{4.5}$$

for infinitesimal ω, then the condition for a Lorentz transformation $\Lambda^\mu{}_\sigma \Lambda^\nu{}_\rho \eta^{\sigma\rho} = \eta^{\mu\nu}$ becomes the requirement that ω is anti-symmetric:

$$\omega^{\mu\nu} + \omega^{\nu\mu} = 0 \tag{4.6}$$

Note that an antisymmetric 4×4 matrix has $4 \times 3/2 = 6$ independent components, which agrees with the 6 transformations of the Lorentz group: 3 rotations and 3 boosts. It's going to be useful to introduce a basis of these six 4×4 anti-symmetric matrices. We could call them $(\mathcal{M}^A)^{\mu\nu}$, with $A = 1, \ldots, 6$. But in fact it's better for us (although initially a little confusing) to replace the single index A with a pair of antisymmetric indices $[\rho\sigma]$, where $\rho, \sigma = 0, \ldots, 3$, so we call our matrices $(\mathcal{M}^{\rho\sigma})^{\mu}{}_{\nu}$. The antisymmetry on the ρ and σ indices means that, for example, $\mathcal{M}^{01} = -\mathcal{M}^{10}$, etc, so that ρ and σ again label six different matrices. Of course, the matrices are also antisymmetric on the $\mu\nu$ indices because they are, after all, antisymmetric matrices. With this notation in place, we can write a basis of six 4×4 antisymmetric matrices as

$$(\mathcal{M}^{\rho\sigma})^{\mu\nu} = \eta^{\rho\mu}\,\eta^{\sigma\nu} - \eta^{\sigma\mu}\eta^{\rho\nu} \tag{4.7}$$

where the indices μ and ν are those of the 4×4 matrix, while ρ and σ denote which basis element we're dealing with. If we use these matrices for anything practical (for example, if we want to multiply them together, or act on some field) we will typically need to lower one index, so we have

$$(\mathcal{M}^{\rho\sigma})^{\mu}{}_{\nu} = \eta^{\rho\mu}\,\delta^{\sigma}{}_{\nu} - \eta^{\sigma\mu}\delta^{\rho}{}_{\nu} \tag{4.8}$$

Since we lowered the index with the Minkowski metric, we pick up various minus signs which means that when written in this form, the matrices are no longer necessarily antisymmetric. Two examples of these basis matrices are,

$$(\mathcal{M}^{01})^{\mu}{}_{\nu} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \quad \text{and} \quad (\mathcal{M}^{12})^{\mu}{}_{\nu} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \tag{4.9}$$

The first, $\mathcal{M}^{01}$, generates boosts in the x^1 direction. It is real and symmetric. The second, $\mathcal{M}^{12}$, generates rotations in the (x^1, x^2)-plane. It is real and antisymmetric. We can now write any ω^{μ}_{ν} as a linear combination of the $\mathcal{M}^{\rho\sigma}$,

$$\omega^{\mu}{}_{\nu} = \tfrac{1}{2}\,\Omega_{\rho\sigma}\,(\mathcal{M}^{\rho\sigma})^{\mu}{}_{\nu} \tag{4.10}$$

where $\Omega_{\rho\sigma}$ are just six numbers (again antisymmetric in the indices) that tell us what Lorentz transformation we're doing. The six basis matrices $\mathcal{M}^{\rho\sigma}$ are called the *generators* of the Lorentz transformations. The generators obey the Lorentz Lie algebra relations,

$$[\mathcal{M}^{\rho\sigma}, \mathcal{M}^{\tau\nu}] = \eta^{\sigma\tau}\mathcal{M}^{\rho\nu} - \eta^{\rho\tau}\mathcal{M}^{\sigma\nu} + \eta^{\rho\nu}\mathcal{M}^{\sigma\tau} - \eta^{\sigma\nu}\mathcal{M}^{\rho\tau} \tag{4.11}$$

where we have suppressed the matrix indices. A finite Lorentz transformation can then be expressed as the exponential

$$\Lambda = \exp\left(\tfrac{1}{2}\,\Omega_{\rho\sigma}\mathcal{M}^{\rho\sigma}\right) \tag{4.12}$$

Let me stress again what each of these objects are: the $\mathcal{M}^{\rho\sigma}$ are six 4×4 basis elements of the Lorentz group; the $\Omega_{\rho\sigma}$ are six numbers telling us what kind of Lorentz transformation we're doing (for example, they say things like rotate by $\theta = \pi/7$ about the x^3-direction and run at speed $v = 0.2$ in the x^1 direction).

4.1 The Spinor Representation

We're interested in finding other matrices which satisfy the Lorentz algebra commutation relations (4.11). We will construct the spinor representation. To do this, we start by defining something which, at first sight, has nothing to do with the Lorentz group. It is the *Clifford algebra*,

$$\{\gamma^\mu, \gamma^\nu\} \equiv \gamma^\mu\gamma^\nu + \gamma^\nu\gamma^\mu = 2\eta^{\mu\nu}1 \tag{4.13}$$

where γ^μ, with $\mu = 0, 1, 2, 3$, are a set of four matrices and the 1 on the right-hand side denotes the unit matrix. This means that we must find four matrices such that

$$\gamma^\mu\gamma^\nu = -\gamma^\nu\gamma^\mu \qquad \text{when } \mu \neq \nu \tag{4.14}$$

and

$$(\gamma^0)^2 = 1 \quad , \quad (\gamma^i)^2 = -1 \qquad i = 1, 2, 3 \tag{4.15}$$

It's not hard to convince yourself that there are no representations of the Clifford algebra using 2×2 or 3×3 matrices. The simplest representation of the Clifford algebra is in terms of 4×4 matrices. There are many such examples of 4×4 matrices which obey (4.13). For example, we may take

$$\gamma^0 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad , \quad \gamma^i = \begin{pmatrix} 0 & \sigma^i \\ -\sigma^i & 0 \end{pmatrix} \tag{4.16}$$

where each element is itself a 2×2 matrix, with the σ^i the Pauli matrices

$$\sigma^1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad , \quad \sigma^2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \quad , \quad \sigma^3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{4.17}$$

which themselves satisfy $\{\sigma^i, \sigma^j\} = 2\delta^{ij}$.

One can construct many other representations of the Clifford algebra by taking $V\gamma^\mu V^{-1}$ for any invertible matrix V. However, up to this equivalence, it turns out that there is a unique irreducible representation of the Clifford algebra. The matrices (4.16) provide one example, known as the *Weyl* or *chiral representation* (for reasons that will soon become clear). We will soon restrict ourselves further, and consider only representations of the Clifford algebra that are related to the chiral representation by a unitary transformation V.

So what does the Clifford algebra have to do with the Lorentz group? Consider the commutator of two γ^μ,

$$S^{\rho\sigma} = \frac{1}{4}\left[\gamma^\rho, \gamma^\sigma\right] = \left\{ \begin{array}{cc} 0 & \rho = \sigma \\ \frac{1}{2}\,\gamma^\rho\gamma^\sigma & \rho \neq \sigma \end{array} \right\} = \tfrac{1}{2}\gamma^\rho\gamma^\sigma - \frac{1}{2}\eta^{\rho\sigma} \tag{4.18}$$

Let's see what properties these matrices have:

Claim 4.1: $[S^{\mu\nu}, \gamma^\rho] = \gamma^\mu\eta^{\nu\rho} - \gamma^\nu\eta^{\rho\mu}$

Proof: When $\mu \neq \nu$ we have

$$\begin{aligned} [S^{\mu\nu}, \gamma^\rho] &= \tfrac{1}{2}[\gamma^\mu\gamma^\nu, \gamma^\rho] \\ &= \tfrac{1}{2}\gamma^\mu\gamma^\nu\gamma^\rho - \tfrac{1}{2}\gamma^\rho\gamma^\mu\gamma^\nu \\ &= \tfrac{1}{2}\gamma^\mu\{\gamma^\nu, \gamma^\rho\} - \tfrac{1}{2}\gamma^\mu\gamma^\rho\gamma^\nu - \tfrac{1}{2}\{\gamma^\rho, \gamma^\mu\}\gamma^\nu + \tfrac{1}{2}\gamma^\mu\gamma^\rho\gamma^\nu \\ &= \gamma^\mu\eta^{\nu\rho} - \gamma^\nu\eta^{\rho\mu} \end{aligned}$$

□

Claim 4.2: The matrices $S^{\mu\nu}$ form a representation of the Lorentz algebra (4.11), meaning

$$[S^{\mu\nu}, S^{\rho\sigma}] = \eta^{\nu\rho}S^{\mu\sigma} - \eta^{\mu\rho}S^{\nu\sigma} + \eta^{\mu\sigma}S^{\nu\rho} - \eta^{\nu\sigma}S^{\mu\rho} \tag{4.19}$$

Proof: Taking $\rho \neq \sigma$, and using Claim 4.1 above, we have

$$\begin{aligned} [S^{\mu\nu}, S^{\rho\sigma}] &= \tfrac{1}{2}[S^{\mu\nu}, \gamma^\rho\gamma^\sigma] \\ &= \tfrac{1}{2}[S^{\mu\nu}, \gamma^\rho]\gamma^\sigma + \tfrac{1}{2}\gamma^\rho[S^{\mu\nu}, \gamma^\sigma] \\ &= \tfrac{1}{2}\gamma^\mu\gamma^\sigma\eta^{\nu\rho} - \tfrac{1}{2}\gamma^\nu\gamma^\sigma\eta^{\rho\mu} + \tfrac{1}{2}\gamma^\rho\gamma^\mu\eta^{\nu\sigma} - \tfrac{1}{2}\gamma^\rho\gamma^\nu\eta^{\sigma\mu} \end{aligned} \tag{4.20}$$

Now using the expression (4.18) to write $\gamma^\mu\gamma^\sigma = 2S^{\mu\sigma} + \eta^{\mu\sigma}$, we have

$$[S^{\mu\nu}, S^{\rho\sigma}] = S^{\mu\sigma}\eta^{\nu\rho} - S^{\nu\sigma}\eta^{\rho\mu} + S^{\rho\mu}\eta^{\nu\sigma} - S^{\rho\nu}\eta^{\sigma\mu} \tag{4.21}$$

which is our desired expression. □

4.1.1 Spinors

The $S^{\mu\nu}$ are 4×4 matrices, because the γ^μ are 4×4 matrices. So far we haven't given an index name to the rows and columns of these matrices: we're going to call them $\alpha, \beta = 1, 2, 3, 4$.

We need a field for the matrices $(S^{\mu\nu})^\alpha{}_\beta$ to act upon. We introduce the Dirac *spinor* field $\psi^\alpha(x)$, an object with four complex components labelled by $\alpha = 1, 2, 3, 4$. Under Lorentz transformations, we have

$$\psi^\alpha(x) \rightarrow S[\Lambda]^\alpha{}_\beta \, \psi^\beta(\Lambda^{-1}x) \tag{4.22}$$

where

$$\Lambda = \exp\left(\tfrac{1}{2}\,\Omega_{\rho\sigma}\mathcal{M}^{\rho\sigma}\right) \tag{4.23}$$

$$S[\Lambda] = \exp\left(\tfrac{1}{2}\,\Omega_{\rho\sigma}S^{\rho\sigma}\right) \tag{4.24}$$

Although the basis of generators $\mathcal{M}^{\rho\sigma}$ and $S^{\rho\sigma}$ are different, we use the same six numbers $\Omega_{\rho\sigma}$ in both Λ and $S[\Lambda]$: this ensures that we're doing the same Lorentz transformation on x and ψ. Note that we denote both the generator $S^{\rho\sigma}$ and the full Lorentz transformation $S[\Lambda]$ as "S". To avoid confusion, the latter will always come with the square brackets $[\Lambda]$.

Both Λ and $S[\Lambda]$ are 4×4 matrices. So how can we be sure that the spinor representation is something new, and isn't equivalent to the familiar representation $\Lambda^\mu{}_\nu$? To see that the two representations are truly different, let's look at some specific transformations.

Rotations

$$S^{ij} = \frac{1}{2}\begin{pmatrix} 0 & \sigma^i \\ -\sigma^i & 0 \end{pmatrix}\begin{pmatrix} 0 & \sigma^j \\ -\sigma^j & 0 \end{pmatrix} = -\frac{i}{2}\,\epsilon^{ijk}\begin{pmatrix} \sigma^k & 0 \\ 0 & \sigma^k \end{pmatrix} \tag{4.25}$$

If we write the rotation parameters as $\Omega_{ij} = -\epsilon_{ijk}\varphi^k$ (meaning $\Omega_{12} = -\varphi^3$, etc) then the rotation matrix becomes

$$S[\Lambda] = \exp\left(\tfrac{1}{2}\,\Omega_{\rho\sigma}S^{\rho\sigma}\right) = \begin{pmatrix} e^{+i\,\vec{\varphi}\cdot\vec{\sigma}/2} & 0 \\ 0 & e^{+i\vec{\varphi}\cdot\vec{\sigma}/2} \end{pmatrix} \tag{4.26}$$

where we need to remember that $\Omega_{12} = -\Omega_{21} = -\varphi^3$ when following factors of 2. Consider now a rotation by 2π about, say, the x^3-axis. This is achieved by $\vec{\varphi} = (0, 0, 2\pi)$,

and the spinor rotation matrix becomes,

$$S[\Lambda] = \begin{pmatrix} e^{+i\pi\sigma^3} & 0 \\ 0 & e^{+i\pi\sigma^3} \end{pmatrix} = -1 \tag{4.27}$$

Therefore under a 2π rotation

$$\psi^\alpha(x) \to -\psi^\alpha(x) \tag{4.28}$$

which is definitely not what happens to a vector! To check that we haven't been cheating with factors of 2, let's see how a vector would transform under a rotation by $\vec{\varphi} = (0, 0, \varphi^3)$. We have

$$\Lambda = \exp\left(\tfrac{1}{2}\,\Omega_{\rho\sigma}\mathcal{M}^{\rho\sigma}\right) = \exp\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & \varphi^3 & 0 \\ 0 & -\varphi^3 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \tag{4.29}$$

So when we rotate a vector by $\varphi^3 = 2\pi$, we learn that $\Lambda = 1$ as you would expect. So $S[\Lambda]$ is definitely a different representation from the familiar vector representation $\Lambda^\mu{}_\nu$.

<u>Boosts</u>

$$S^{0i} = \frac{1}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 0 & \sigma^i \\ -\sigma^i & 0 \end{pmatrix} = \frac{1}{2}\begin{pmatrix} -\sigma^i & 0 \\ 0 & \sigma^i \end{pmatrix} \tag{4.30}$$

Writing the boost parameter as $\Omega_{i0} = -\Omega_{0i} = \chi_i$, we have

$$S[\Lambda] = \begin{pmatrix} e^{+\vec{\chi}\cdot\vec{\sigma}/2} & 0 \\ 0 & e^{-\vec{\chi}\cdot\vec{\sigma}/2} \end{pmatrix} \tag{4.31}$$

Representations of the Lorentz Group are not Unitary

Note that for rotations given in (4.26), $S[\Lambda]$ is unitary, satisfying $S[\Lambda]^\dagger S[\Lambda] = 1$. But for boosts given in (4.31), $S[\Lambda]$ is not unitary. In fact, there are *no* finite dimensional unitary representations of the Lorentz group. We have demonstrated this explicitly for the spinor representation using the chiral representation (4.16) of the Clifford algebra. We can get a feel for why it is true for a spinor representation constructed from any representation of the Clifford algebra. Recall that

$$S[\Lambda] = \exp\left(\tfrac{1}{2}\,\Omega_{\rho\sigma}S^{\rho\sigma}\right) \tag{4.32}$$

so the representation is unitary if $S^{\mu\nu}$ are anti-hermitian, i.e. $(S^{\mu\nu})^\dagger = -S^{\mu\nu}$. But we have

$$(S^{\mu\nu})^\dagger = -\frac{1}{4}[(\gamma^\mu)^\dagger, (\gamma^\nu)^\dagger] \tag{4.33}$$

which can be anti-hermitian if all γ^μ are hermitian or all are anti-hermitian. However, we can never arrange for this to happen since

$$\begin{aligned} (\gamma^0)^2 = 1 &\quad \Rightarrow \quad \text{Real Eigenvalues} \\ (\gamma^i)^2 = -1 &\quad \Rightarrow \quad \text{Imaginary Eigenvalues} \end{aligned} \tag{4.34}$$

So we could pick γ^0 to be hermitian, but we can only pick γ^i to be anti-hermitian. Indeed, in the chiral representation (4.16), the matrices have this property: $(\gamma^0)^\dagger = \gamma^0$ and $(\gamma^i)^\dagger = -\gamma^i$. In general there is no way to pick γ^μ such that $S^{\mu\nu}$ are anti-hermitian.

4.2 Constructing an Action

We now have a new field to work with, the Dirac spinor ψ. We would like to construct a Lorentz invariant equation of motion. We do this by constructing a Lorentz invariant action.

We will start in a naive way which won't work, but will give us a clue how to proceed. Define

$$\psi^\dagger(x) = (\psi^\star)^T(x) \tag{4.35}$$

which is the usual adjoint of a multi-component object. We could then try to form a Lorentz scalar by taking the product $\psi^\dagger\psi$, with the spinor indices summed over. Let's see how this transforms under Lorentz transformations,

$$\begin{aligned} \psi(x) &\to S[\Lambda]\,\psi(\Lambda^1 x) \\ \psi^\dagger(x) &\to \psi^\dagger(\Lambda^{-1}x)\,S[\Lambda]^\dagger \end{aligned} \tag{4.36}$$

So $\psi^\dagger(x)\psi(x) \to \psi^\dagger(\Lambda^{-1}x)S[\Lambda]^\dagger S[\Lambda]\psi(\Lambda^{-1}x)$. But, as we have seen, for some Lorentz transformation $S[\Lambda]^\dagger S[\Lambda] \neq 1$ since the representation is not unitary. This means that $\psi^\dagger\psi$ isn't going to do it for us: it doesn't have any nice transformation under the Lorentz group, and certainly isn't a scalar. But now we see why it fails, we can also see how to proceed. Let's pick a representation of the Clifford algebra which, like the chiral representation (4.16), satisfies $(\gamma^0)^\dagger = \gamma^0$ and $(\gamma^i)^\dagger = -\gamma^i$. Then for all $\mu = 0, 1, 2, 3$ we have

$$\gamma^0\gamma^\mu\gamma^0 = (\gamma^\mu)^\dagger \tag{4.37}$$

which, in turn, means that

$$(S^{\mu\nu})^\dagger = \frac{1}{4}[(\gamma^\nu)^\dagger, (\gamma^\mu)^\dagger] = -\gamma^0 S^{\mu\nu}\gamma^0 \tag{4.38}$$

so that

$$S[\Lambda]^\dagger = \exp\left(\tfrac{1}{2}\,\Omega_{\rho\sigma}(S^{\rho\sigma})^\dagger\right) = \gamma^0\,S[\Lambda]^{-1}\gamma^0 \tag{4.39}$$

With this in mind, we now define the *Dirac adjoint*

$$\bar\psi(x) = \psi^\dagger(x)\,\gamma^0 \tag{4.40}$$

Let's now see what Lorentz covariant objects we can form out of a Dirac spinor ψ and its adjoint $\bar\psi$.

Claim 4.3: $\bar\psi\psi$ is a Lorentz scalar.

Proof: Under a Lorentz transformation,

$$\begin{aligned}\bar\psi(x)\,\psi(x) &= \psi^\dagger(x)\,\gamma^0\,\psi(x)\\ &\to \psi^\dagger(\Lambda^{-1}x)\,S[\Lambda]^\dagger\gamma^0\,S[\Lambda]\psi(\Lambda^{-1}x)\\ &= \psi^\dagger(\Lambda^{-1}x)\,\gamma^0\,\psi(\Lambda^{-1}x)\\ &= \bar\psi(\Lambda^{-1}x)\,\psi(\Lambda^{-1}x)\end{aligned} \tag{4.41}$$

which is indeed the transformation law for a Lorentz scalar. □

Claim 4.4: $\bar\psi\,\gamma^\mu\,\psi$ is a Lorentz vector, which means that

$$\bar\psi(x)\,\gamma^\mu\,\psi(x) \to \Lambda^\mu{}_\nu\,\bar\psi(\Lambda^{-1}x)\,\gamma^\nu\,\psi(\Lambda^{-1}x) \tag{4.42}$$

This equation means that we can treat the $\mu = 0,1,2,3$ index on the γ^μ matrices as a true vector index. In particular we can form Lorentz scalars by contracting it with other Lorentz indices.

Proof: Suppressing the x argument, under a Lorentz transformation we have,

$$\bar\psi\,\gamma^\mu\,\psi \to \bar\psi\,S[\Lambda]^{-1}\,\gamma^\mu\,S[\Lambda]\psi \tag{4.43}$$

If $\bar\psi\,\gamma^\mu\,\psi$ is to transform as a vector, we must have

$$S[\Lambda]^{-1}\,\gamma^\mu\,S[\Lambda] = \Lambda^\mu{}_\nu\gamma^\nu \tag{4.44}$$

We'll now show this. We work infinitesimally, so that

$$\Lambda = \exp\left(\tfrac{1}{2}\,\Omega_{\rho\sigma}\mathcal{M}^{\rho\sigma}\right) \approx 1 + \tfrac{1}{2}\Omega_{\rho\sigma}\mathcal{M}^{\rho\sigma} + \ldots \tag{4.45}$$

$$S[\Lambda] = \exp\left(\tfrac{1}{2}\,\Omega_{\rho\sigma}S^{\rho\sigma}\right) \approx 1 + \tfrac{1}{2}\Omega_{\rho\sigma}S^{\rho\sigma} + \ldots \tag{4.46}$$

so the requirement (4.44) becomes

$$-[S^{\rho\sigma}, \gamma^\mu] = (\mathcal{M}^{\rho\sigma})^\mu{}_\nu \gamma^\nu \tag{4.47}$$

where we've suppressed the α, β indices on γ^μ and $S^{\mu\nu}$, but otherwise left all other indices explicit. In fact equation (4.47) follows from Claim 4.1 where we showed that $[S^{\rho\sigma}, \gamma^\mu] = \gamma^\rho \eta^{\sigma\mu} - \gamma^\sigma \eta^{\mu\rho}$. To see this, we write the right-hand side of (4.47) by expanding out $\mathcal{M}$,

$$\begin{aligned}(\mathcal{M}^{\rho\sigma})^\mu{}_\nu \gamma^\nu &= (\eta^{\rho\mu}\delta^\sigma_\nu - \eta^{\sigma\mu}\delta^\rho_\nu)\gamma^\nu \\ &= \eta^{\rho\mu}\gamma^\sigma - \eta^{\sigma\mu}\gamma^\rho \end{aligned} \tag{4.48}$$

which means that the proof follows if we can show

$$-[S^{\rho\sigma}, \gamma^\mu] = \eta^{\rho\mu}\gamma^\sigma - \eta^{\sigma\mu}\gamma^\rho \tag{4.49}$$

which is exactly what we proved in Claim 4.1. □

Claim 4.5: $\bar{\psi}\gamma^\mu\gamma^\nu\psi$ transforms as a Lorentz tensor. More precisely, the symmetric part is a Lorentz scalar, proportional to $\eta^{\mu\nu}\,\bar{\psi}\psi$, while the antisymmetric part is a Lorentz tensor, proportional to $\bar{\psi}S^{\mu\nu}\psi$.

Proof: As above. □

We are now armed with three bilinears of the Dirac field, $\bar{\psi}\psi$, $\bar{\psi}\gamma^\mu\psi$ and $\bar{\psi}\gamma^\mu\gamma^\nu\psi$, each of which transforms covariantly under the Lorentz group. We can try to build a Lorentz invariant action from these. In fact, we need only the first two. We choose

$$S = \int d^4x\; \bar{\psi}(x)\; (i\gamma^\mu\partial_\mu - m)\; \psi(x) \tag{4.50}$$

This is the Dirac action. The factor of "i" is there to make the action real; upon complex conjugation, it cancels a minus sign that comes from integration by parts. (Said another way, it's there for the same reason that the Hermitian momentum operator $-i\nabla$ in quantum mechanics has a factor i). As we will see in the next section, after quantization this theory describes particles and anti-particles of mass $|m|$ and spin 1/2. Notice that the Lagrangian is first order, rather than the second order Lagrangians we were working with for scalar fields. Also, the mass appears in the Lagrangian as m, which can be positive or negative.

4.3 The Dirac Equation

The equation of motion follows from the action (4.50) by varying with respect to ψ and $\bar{\psi}$ independently. Varying with respect to $\bar{\psi}$, we have

$$(i\gamma^\mu \partial_\mu - m)\,\psi = 0 \tag{4.51}$$

This is the *Dirac equation.* It's completely gorgeous. Varying with respect to ψ gives the conjugate equation

$$i\partial_\mu \bar{\psi}\,\gamma^\mu + m\bar{\psi} = 0 \tag{4.52}$$

The Dirac equation is first order in derivatives, yet miraculously Lorentz invariant. If we tried to write down a first order equation of motion for a scalar field, it would look like $v^\mu \partial_\mu \phi = \ldots$, which necessarily includes a privileged vector in spacetime v^μ and is not Lorentz invariant. However, for spinor fields, the magic of the γ^μ matrices means that the Dirac Lagrangian is Lorentz invariant.

The Dirac equation mixes up different components of ψ through the matrices γ^μ. However, each individual component itself solves the Klein-Gordon equation. To see this, write

$$(i\gamma^\nu \partial_\nu + m)(i\gamma^\mu \partial_\mu - m)\,\psi = -\left(\gamma^\mu \gamma^\nu \partial_\mu \partial_\nu + m^2\right)\psi = 0 \tag{4.53}$$

But $\gamma^\mu \gamma^\nu \partial_\mu \partial_\nu = \frac{1}{2}\{\gamma^\mu, \gamma^\nu\}\partial_\mu \partial_\nu = \partial_\mu \partial^\mu$, so we get

$$-(\partial_\mu \partial^\mu + m^2)\psi = 0 \tag{4.54}$$

where this last equation has no γ^μ matrices, and so applies to each component ψ^α, with $\alpha = 1, 2, 3, 4$.

The Slash

Let's introduce some useful notation. We will often come across 4-vectors contracted with γ^μ matrices. We write

$$A_\mu \gamma^\mu \equiv \not{A} \tag{4.55}$$

so the Dirac equation reads

$$(i\not{\partial} - m)\psi = 0 \tag{4.56}$$

4.4 Chiral Spinors

When we've needed an explicit form of the γ^μ matrices, we've used the chiral representation

$$\gamma^0 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad , \quad \gamma^i = \begin{pmatrix} 0 & \sigma^i \\ -\sigma^i & 0 \end{pmatrix} \tag{4.57}$$

In this representation, the spinor rotation transformation $S[\Lambda_{\rm rot}]$ and boost transformation $S[\Lambda_{\rm boost}]$ were computed in (4.26) and (4.31). Both are block diagonal,

$$S[\Lambda_{\rm rot}] = \begin{pmatrix} e^{+i\,\vec{\varphi}\cdot\vec{\sigma}/2} & 0 \\ 0 & e^{+i\vec{\varphi}\cdot\vec{\sigma}/2} \end{pmatrix} \quad \text{and} \quad S[\Lambda_{\rm boost}] = \begin{pmatrix} e^{+\vec{\chi}\cdot\vec{\sigma}/2} & 0 \\ 0 & e^{-\vec{\chi}\cdot\vec{\sigma}/2} \end{pmatrix} \tag{4.58}$$

This means that the Dirac spinor representation of the Lorentz group is *reducible*. It decomposes into two irreducible representations, acting only on two-component spinors $u_\pm$ which, in the chiral representation, are defined by

$$\psi = \begin{pmatrix} u_+ \\ u_- \end{pmatrix} \tag{4.59}$$

The two-component objects $u_\pm$ are called *Weyl spinors* or *chiral spinors*. They transform in the same way under rotations,

$$u_\pm \rightarrow e^{i\vec{\varphi}\cdot\vec{\sigma}/2}\, u_\pm \tag{4.60}$$

but oppositely under boosts,

$$u_\pm \rightarrow e^{\pm\vec{\chi}\cdot\vec{\sigma}/2} u_\pm \tag{4.61}$$

In group theory language, u_+ is in the $(\frac{1}{2},0)$ representation of the Lorentz group, while u_- is in the $(0,\frac{1}{2})$ representation. The Dirac spinor ψ lies in the $(\frac{1}{2},0)\oplus(0,\frac{1}{2})$ representation. (Strictly speaking, the spinor is a representation of the double cover of the Lorentz group $SL(2,\mathbf{C})$).

4.4.1 The Weyl Equation

Let's see what becomes of the Dirac Lagrangian under the decomposition (4.59) into Weyl spinors. We have

$$\mathcal{L} = \bar{\psi}(i\,\partial\!\!\!/ - m)\psi = iu_-^\dagger \sigma^\mu \partial_\mu u_- + iu_+^\dagger \bar{\sigma}^\mu \partial_\mu u_+ - m(u_+^\dagger u_- + u_-^\dagger u_+) = 0 \tag{4.62}$$

where we have introduced some new notation for the Pauli matrices with a $\mu = 0, 1, 2, 3$ index,

$$\sigma^\mu = (1, \sigma^i) \quad \text{and} \quad \bar{\sigma}^\mu = (1, -\sigma^i) \tag{4.63}$$

From (4.62), we see that a massive fermion requires both u_+ and u_-, since they couple through the mass term. However, a massless fermion can be described by u_+ (or u_-) alone, with the equation of motion

$$\begin{aligned} & i\bar{\sigma}^\mu \partial_\mu u_+ = 0 \\ \text{or} \quad & i\sigma^\mu \partial_\mu u_- = 0 \end{aligned} \tag{4.64}$$

These are the *Weyl equations.*

Degrees of Freedom

Let me comment here on the degrees of freedom in a spinor. The Dirac fermion has 4 complex components = 8 real components. How do we count degrees of freedom? In classical mechanics, the number of degrees of freedom of a system is equal to the dimension of the configuration space or, equivalently, half the dimension of the phase space. In field theory we have an infinite number of degrees of freedom, but it makes sense to count the number of degrees of freedom per spatial point: this should at least be finite. For example, in this sense a real scalar field ϕ has a single degree of freedom. At the quantum level, this translates to the fact that it gives rise to a single type of particle. A classical complex scalar field has two degrees of freedom, corresponding to the particle and the anti-particle in the quantum theory.

But what about a Dirac spinor? One might think that there are 8 degrees of freedom. But this isn't right. Crucially, and in contrast to the scalar field, the equation of motion is first order rather than second order. In particular, for the Dirac Lagrangian, the momentum conjugate to the spinor ψ is given by

$$\pi_\psi = \partial \mathcal{L} / \partial \dot{\psi} = i\psi^\dagger \tag{4.65}$$

It is not proportional to the time derivative of ψ. This means that the phase space for a spinor is therefore parameterized by ψ and $\psi^\dagger$, while for a scalar it is parameterized by ϕ and $\pi = \dot{\phi}$. So the *phase space* of the Dirac spinor ψ has 8 real dimensions and correspondingly the number of real degrees of freedom is 4. We will see in the next section that, in the quantum theory, this counting manifests itself as two degrees of freedom (spin up and down) for the particle, and a further two for the anti-particle.

A similar counting for the Weyl fermion tells us that it has two degrees of freedom.

4.4.2 γ^5

The Lorentz group matrices $S[\Lambda]$ came out to be block diagonal in (4.58) because we chose the specific representation (4.57). In fact, this is why the representation (4.57) is called the chiral representation: it's because the decomposition of the Dirac spinor ψ is simply given by (4.59). But what happens if we choose a different representation γ^μ of the Clifford algebra, so that

$$\gamma^\mu \to U\gamma^\mu U^{-1} \quad \text{and} \quad \psi \to U\psi \ \ ? \tag{4.66}$$

Now $S[\Lambda]$ will not be block diagonal. Is there an invariant way to define chiral spinors? We can do this by introducing the "fifth" gamma-matrix

$$\gamma^5 = -i\gamma^0\gamma^1\gamma^2\gamma^3 \tag{4.67}$$

You can check that this matrix satisfies

$$\{\gamma^5, \gamma^\mu\} = 0 \quad \text{and} \quad (\gamma^5)^2 = +1 \tag{4.68}$$

The reason that this is called γ^5 is because the set of matrices $\tilde{\gamma}^A = (\gamma^\mu, i\gamma^5)$, with $A = 0, 1, 2, 3, 4$ satisfy the five-dimensional Clifford algebra $\{\tilde{\gamma}^A, \tilde{\gamma}^B\} = 2\eta^{AB}$. (You might think that γ^4 would be a better name! But γ^5 is the one everyone chooses - it's a more sensible name in Euclidean space, where $A = 1, 2, 3, 4, 5$). You can also check that $[S_{\mu\nu}, \gamma^5] = 0$, which means that γ^5 is a scalar under rotations and boosts. Since $(\gamma^5)^2 = 1$, this means we may form the Lorentz invariant projection operators

$$P_\pm = \frac{1}{2}\left(1 \pm \gamma^5\right) \tag{4.69}$$

such that $P_+^2 = P_+$ and $P_-^2 = P_-$ and $P_+P_- = 0$. One can check that for the chiral representation (4.57),

$$\gamma^5 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{4.70}$$

from which we see that the operators $P_\pm$ project onto the Weyl spinors $u_\pm$. However, for an arbitrary representation of the Clifford algebra, we may use γ^5 to define the chiral spinors,

$$\psi_\pm = P_\pm \psi \tag{4.71}$$

which form the irreducible representations of the Lorentz group. ψ_+ is often called a "right-handed" spinor, while ψ_- is "left-handed".

4.4.3 Parity

The spinors $\psi_\pm$ are related to each other by *parity*. Let's pause to define this concept. The Lorentz group is defined by $x^\mu \to \Lambda^\mu{}_\nu x^\nu$ such that

$$\Lambda^\mu{}_\nu \Lambda^\rho{}_\sigma \eta^{\nu\sigma} = \eta^{\mu\rho} \tag{4.72}$$

So far we have only considered transformations Λ which are continuously connected to the identity; these are the ones which have an infinitesimal form. However there are also two discrete symmetries which are part of the Lorentz group. They are

$$\begin{aligned} \text{Time Reversal} \quad & T : x^0 \to -x^0 \ ; \ x^i \to x^i \\ \text{Parity} \quad & P : x^0 \to x^0 \ ; \ x^i \to -x^i \end{aligned} \tag{4.73}$$

We won't discuss time reversal too much in this course. (It turns out to be represented by an anti-unitary transformation on states. See, for example the book by Peskin and Schroeder). But parity has an important role to play in the standard model and, in particular, the theory of the weak interaction.

Under parity, the left and right-handed spinors are exchanged. This follows from the transformation of the spinors under the Lorentz group. In the chiral representation, we saw that the rotation (4.60) and boost (4.61) transformations for the Weyl spinors $u_\pm$ are

$$u_\pm \xrightarrow{\text{rot}} e^{i\vec{\varphi}\cdot\vec{\sigma}/2}\, u_\pm \quad \text{and} \quad u_\pm \xrightarrow{\text{boost}} e^{\pm\vec{\chi}\cdot\vec{\sigma}/2} u_\pm \tag{4.74}$$

Under parity, rotations don't change sign. But boosts do flip sign. This confirms that parity exchanges right-handed and left-handed spinors, $P : u_\pm \to u_\mp$, or in the notation $\psi_\pm = \frac{1}{2}(1 \pm \gamma^5)\psi$, we have

$$P : \psi_\pm(\vec{x}, t) \to \psi_\mp(-\vec{x}, t) \tag{4.75}$$

Using this knowledge of how chiral spinors transform, and the fact that $P^2 = 1$, we see that the action of parity on the Dirac spinor itself can be written as

$$P : \ \psi(\vec{x}, t) \to \gamma^0 \psi(-\vec{x}, t) \tag{4.76}$$

Notice that if $\psi(\vec{x}, t)$ satisfies the Dirac equation, then the parity transformed spinor $\gamma^0\psi(-\vec{x}, t)$ also satisfies the Dirac equation, meaning

$$(i\gamma^0\partial_t + i\gamma^i\partial_i - m)\gamma^0\psi(-\vec{x}, t) = \gamma^0(i\gamma^0\partial_t - i\gamma^i\partial_i - m)\psi(-\vec{x}, t) = 0 \tag{4.77}$$

where the extra minus sign from passing γ^0 through γ^i is compensated by the derivative acting on $-\vec{x}$ instead of $+\vec{x}$.

4.4.4 Chiral Interactions

Let's now look at how our interaction terms change under parity. We can look at each of our spinor bilinears from which we built the action,

$$P: \ \bar{\psi}\psi(\vec{x},t) \rightarrow \bar{\psi}\psi(-\vec{x},t) \tag{4.78}$$

which is the transformation of a scalar. For the vector $\bar{\psi}\gamma^\mu\psi$, we can look at the temporal and spatial components separately,

$$\begin{aligned} P: \ \bar{\psi}\gamma^0\psi(\vec{x},t) &\rightarrow \bar{\psi}\gamma^0\psi(-\vec{x},t) \\ P: \ \bar{\psi}\gamma^i\psi(\vec{x},t) &\rightarrow \bar{\psi}\gamma^0\gamma^i\gamma^0\psi(-\vec{x},t) = -\bar{\psi}\gamma^i\psi(-\vec{x},t) \end{aligned} \tag{4.79}$$

which tells us that $\bar{\psi}\gamma^\mu\psi$ transforms as a vector, with the spatial part changing sign. You can also check that $\bar{\psi}S^{\mu\nu}\psi$ transforms as a suitable tensor.

However, now we've discovered the existence of γ^5, we can form another Lorentz scalar and another Lorentz vector,

$$\bar{\psi}\gamma^5\psi \quad \text{and} \quad \bar{\psi}\gamma^5\gamma^\mu\psi \tag{4.80}$$

How do these transform under parity? We can check:

$$P: \ \bar{\psi}\gamma^5\psi(\vec{x},t) \rightarrow \bar{\psi}\gamma^0\gamma^5\gamma^0 \ \psi(-\vec{x},t) = -\bar{\psi}\gamma^5\psi(-\vec{x},t) \tag{4.81}$$

$$P: \ \bar{\psi}\gamma^5\gamma^\mu\psi(\vec{x},t) \rightarrow \bar{\psi}\gamma^0\gamma^5\gamma^\mu\gamma^0\psi(-\vec{x},t) = \begin{cases} -\bar{\psi}\gamma^5\gamma^0\psi(-\vec{x},t) & \mu = 0 \\ +\bar{\psi}\gamma^5\gamma^i\psi(-\vec{x},t) & \mu = i \end{cases}$$

which means that $\bar{\psi}\gamma^5\psi$ transforms as a *pseudoscalar*, while $\bar{\psi}\gamma^5\gamma^\mu\psi$ transforms as an *axial vector*. To summarize, we have the following spinor bilinears,

$$\begin{aligned} \bar{\psi}\psi &: \quad \text{scalar} \\ \bar{\psi}\gamma^\mu\psi &: \quad \text{vector} \\ \bar{\psi}S^{\mu\nu}\psi &: \quad \text{tensor} \\ \bar{\psi}\gamma^5\psi &: \quad \text{pseudoscalar} \\ \bar{\psi}\gamma^5\gamma^\mu\psi &: \quad \text{axial vector} \end{aligned} \tag{4.82}$$

The total number of bilinears is $1 + 4 + (4 \times 3/2) + 4 + 1 = 16$ which is all we could hope for from a 4-component object.

We're now armed with new terms involving γ^5 that we can start to add to our Lagrangian to construct new theories. Typically such terms will break parity invariance of the theory, although this is not always true. (For example, the term $\phi\bar{\psi}\gamma^5\psi$ doesn't break parity if ϕ is itself a pseudoscalar). Nature makes use of these parity violating interactions by using γ^5 in the weak force. A theory which treats $\psi_\pm$ on an equal footing is called a *vector-like theory*. A theory in which ψ_+ and ψ_- appear differently is called a *chiral theory*.

4.5 Majorana Fermions

Our spinor ψ^α is a complex object. It has to be because the representation $S[\Lambda]$ is typically also complex. This means that if we were to try to make ψ real, for example by imposing $\psi = \psi^\star$, then it wouldn't stay that way once we make a Lorentz transformation. However, there is a way to impose a reality condition on the Dirac spinor ψ. To motivate this possibility, it's simplest to look at a novel basis for the Clifford algebra, known as the *Majorana basis*.

$$\gamma^0 = \begin{pmatrix} 0 & \sigma^2 \\ \sigma^2 & 0 \end{pmatrix} \; , \quad \gamma^1 = \begin{pmatrix} i\sigma^3 & 0 \\ 0 & i\sigma^3 \end{pmatrix} \; , \quad \gamma^2 = \begin{pmatrix} 0 & -\sigma^2 \\ \sigma^2 & 0 \end{pmatrix} \; , \quad \gamma^3 = \begin{pmatrix} -i\sigma^1 & 0 \\ 0 & -i\sigma^1 \end{pmatrix}$$

These matrices satisfy the Clifford algebra. What is special about them is that they are all pure imaginary $(\gamma^\mu)^\star = -\gamma^\mu$. This means that the generators of the Lorentz group $S^{\mu\nu} = \frac{1}{4}[\gamma^\mu, \gamma^\nu]$, and hence the matrices $S[\Lambda]$ are real. So with this basis of the Clifford algebra, we can work with a real spinor simply by imposing the condition,

$$\psi = \psi^\star \tag{4.83}$$

which is preserved under Lorentz transformation. Such spinors are called *Majorana spinors*.

So what's the story if we use a general basis for the Clifford algebra? We'll ask only that the basis satisfies $(\gamma^0)^\dagger = \gamma^0$ and $(\gamma^i)^\dagger = -\gamma^i$. We then define the *charge conjugate* of a Dirac spinor ψ as

$$\psi^{(c)} = C\psi^\star \tag{4.84}$$

Here C is a 4×4 matrix satisfying

$$C^\dagger C = 1 \quad \text{and} \quad C^\dagger \gamma^\mu C = -(\gamma^\mu)^\star \tag{4.85}$$

Let's firstly check that (4.84) is a good definition, meaning that $\psi^{(c)}$ transforms nicely under a Lorentz transformation. We have

$$\psi^{(c)} \to C S[\Lambda]^\star \psi^\star = S[\Lambda] C \psi^\star = S[\Lambda]\psi^{(c)} \tag{4.86}$$

where we've made use of the properties (4.85) in taking the matrix C through $S[\Lambda]^\star$. In fact, not only does $\psi^{(c)}$ transform nicely under the Lorentz group, but if ψ satisfies the Dirac equation, then $\psi^{(c)}$ does too. This follows from,

$$\begin{aligned}(i\,\partial\!\!\!/ - m)\psi = 0 \quad &\Rightarrow \quad (-i\,\partial\!\!\!/^\star - m)\psi^\star = 0 \\ &\Rightarrow \quad C(-i\,\partial\!\!\!/^\star - m)\psi^\star = (+i\,\partial\!\!\!/ - m)\psi^{(c)} = 0\end{aligned}$$

Finally, we can now impose the Lorentz invariant reality condition on the Dirac spinor, to yield a Majorana spinor,

$$\psi^{(c)} = \psi \tag{4.87}$$

After quantization, the Majorana spinor gives rise to a fermion that is its own anti-particle. This is exactly the same as in the case of scalar fields, where we've seen that a real scalar field gives rise to a spin 0 boson that is its own anti-particle. (Be aware: In many texts an extra factor of γ^0 is absorbed into the definition of C).

So what is this matrix C? Well, for a given representation of the Clifford algebra, it is something that we can find fairly easily. In the Majorana basis, where the gamma matrices are pure imaginary, we have simply $C_{\text{Maj}} = 1$ and the Majorana condition $\psi = \psi^{(c)}$ becomes $\psi = \psi^\star$. In the chiral basis (4.16), only γ^2 is imaginary, and we may take $C_{\text{chiral}} = i\gamma^2 = \left(\begin{smallmatrix} 0 & i\sigma^2 \\ -i\sigma^2 & 0 \end{smallmatrix}\right)$. (The matrix $i\sigma^2$ that appears here is simply the anti-symmetric matrix $\epsilon^{\alpha\beta}$). It is interesting to see how the Majorana condition (4.87) looks in terms of the decomposition into left and right handed Weyl spinors (4.59). Plugging in the various definitions, we find that $u_+ = i\sigma^2 u_-^\star$ and $u_- = -i\sigma^2 u_+^\star$. In other words, a Majorana spinor can be written in terms of Weyl spinors as

$$\psi = \begin{pmatrix} u_+ \\ -i\sigma^2 u_+^\star \end{pmatrix} \tag{4.88}$$

Notice that it's not possible to impose the Majorana condition $\psi = \psi^{(c)}$ at the same time as the Weyl condition ($u_- = 0$ or $u_+ = 0$). Instead the Majorana condition relates u_- and u_+.

An Aside: Spinors in Different Dimensions: The ability to impose Majorana or Weyl conditions on Dirac spinors depends on both the dimension and the signature of spacetime. One can always impose the Weyl condition on a spinor in even dimensional Minkowski space, basically because you can always build a suitable "γ^5" projection matrix by multiplying together all the other γ-matrices. The pattern for when the Majorana condition can be imposed is a little more sporadic. Interestingly, although the Majorana condition and Weyl condition cannot be imposed simultaneously in four dimensions, you can do this in Minowski spacetimes of dimension $2, 10, 18, \ldots$.

4.6 Symmetries and Conserved Currents

The Dirac Lagrangian enjoys a number of symmetries. Here we list them and compute the associated conserved currents.

Spacetime Translations

Under spacetime translations the spinor transforms as

$$\delta\psi = \epsilon^\mu \partial_\mu \psi \tag{4.89}$$

The Lagrangian depends on $\partial_\mu\psi$, but not $\partial_\mu\bar{\psi}$, so the standard formula (1.41) gives us the energy-momentum tensor

$$T^{\mu\nu} = i\bar{\psi}\gamma^\mu\partial^\nu\psi - \eta^{\mu\nu}\mathcal{L} \tag{4.90}$$

Since a current is conserved only when the equations of motion are obeyed, we don't lose anything by imposing the equations of motion already on $T^{\mu\nu}$. In the case of a scalar field this didn't really buy us anything because the equations of motion are second order in derivatives, while the energy-momentum is typically first order. However, for a spinor field the equations of motion are first order: $(i\partial\!\!\!/ - m)\psi = 0$. This means we can set $\mathcal{L} = 0$ in $T^{\mu\nu}$, leaving

$$T^{\mu\nu} = i\bar{\psi}\gamma^\mu\partial^\nu\psi \tag{4.91}$$

In particular, we have the total energy

$$E = \int d^3x\, T^{00} = \int d^3x\, i\bar{\psi}\gamma^0\dot{\psi} = \int d^3x\, \psi^\dagger\gamma^0(-i\gamma^i\partial_i + m)\psi \tag{4.92}$$

where, in the last equality, we have again used the equations of motion.

Lorentz Transformations

Under an infinitesimal Lorentz transformation, the Dirac spinor transforms as (4.22) which, in infinitesimal form, reads

$$\delta\psi^\alpha = -\omega^\mu_{\ \nu}x^\nu\partial_\mu\psi^\alpha + \tfrac{1}{2}\Omega_{\rho\sigma}(S^{\rho\sigma})^\alpha_{\ \beta}\psi^\beta \tag{4.93}$$

where, following (4.10), we have $\omega^\mu_{\ \nu} = \frac{1}{2}\Omega_{\rho\sigma}(\mathcal{M}^{\rho\sigma})^\mu_{\ \nu}$, and $\mathcal{M}^{\rho\sigma}$ are the generators of the Lorentz algebra given by (4.8)

$$(\mathcal{M}^{\rho\sigma})^\mu_\nu = \eta^{\rho\mu}\,\delta^\sigma_\nu - \eta^{\sigma\mu}\delta^\rho_\nu \tag{4.94}$$

which, after direct substitution, tells us that $\omega^{\mu\nu} = \Omega^{\mu\nu}$. So we get

$$\delta\psi^\alpha = -\omega^{\mu\nu}\left[x_\nu \partial_\mu \psi^\alpha - \tfrac{1}{2}(S_{\mu\nu})^\alpha{}_\beta \psi^\beta\right] \tag{4.95}$$

The conserved current arising from Lorentz transformations now follows from the same calculation we saw for the scalar field (1.54) with two differences: firstly, as we saw above, the spinor equations of motion set $\mathcal{L} = 0$; secondly, we pick up an extra piece in the current from the second term in (4.95). We have

$$(\mathcal{J}^\mu)^{\rho\sigma} = x^\sigma T^{\mu\rho} - x^\rho T^{\mu\sigma} - i\bar{\psi}\gamma^\mu S^{\rho\sigma}\psi \tag{4.96}$$

After quantization, when $(\mathcal{J}^\mu)^{\rho\sigma}$ is turned into an operator, this extra term will be responsible for providing the single particle states with internal angular momentum, telling us that the quantization of a Dirac spinor gives rise to a particle carrying spin 1/2.

Internal Vector Symmetry

The Dirac Lagrangian is invariant under rotating the phase of the spinor, $\psi \to e^{-i\alpha}\psi$. This gives rise to the current

$$j_V^\mu = \bar{\psi}\gamma^\mu\psi \tag{4.97}$$

where "V" stands for *vector*, reflecting the fact that the left and right-handed components $\psi_\pm$ transform in the same way under this symmetry. We can easily check that j_V^μ is conserved under the equations of motion,

$$\partial_\mu j_V^\mu = (\partial_\mu\bar{\psi})\gamma^\mu\psi + \bar{\psi}\gamma^\mu(\partial_\mu\psi) = im\bar{\psi}\psi - im\bar{\psi}\psi = 0 \tag{4.98}$$

where, in the last equality, we have used the equations of motion $i\partial\!\!\!/\psi = m\psi$ and $i\partial_\mu\bar{\psi}\gamma^\mu = -m\bar{\psi}$. The conserved quantity arising from this symmetry is

$$Q = \int d^3x\, \bar{\psi}\gamma^0\psi = \int d^3x\ \psi^\dagger\psi \tag{4.99}$$

We will see shortly that this has the interpretation of electric charge, or particle number, for fermions.

Axial Symmetry

When $m = 0$, the Dirac Lagrangian admits an extra internal symmetry which rotates left and right-handed fermions in opposite directions,

$$\psi \to e^{i\alpha\gamma^5}\psi \quad \text{and} \quad \bar{\psi} \to \bar{\psi}e^{i\alpha\gamma^5} \tag{4.100}$$

Here the second transformation follows from the first after noting that $e^{-i\alpha\gamma^5}\gamma^0 = \gamma^0 e^{+i\alpha\gamma^5}$. This gives the conserved current,

$$j^\mu_A = \bar{\psi}\gamma^\mu\gamma^5\psi \tag{4.101}$$

where A is for "axial" since j^μ_A is an axial vector. This is conserved only when $m = 0$. Indeed, with the full Dirac Lagrangian we may compute

$$\partial_\mu j^\mu_A = (\partial_\mu\bar{\psi})\gamma^\mu\gamma^5\psi + \bar{\psi}\gamma^\mu\gamma^5\partial_\mu\psi = 2im\bar{\psi}\gamma^5\psi \tag{4.102}$$

which vanishes only for $m = 0$. However, in the quantum theory things become more interesting for the axial current. When the theory is coupled to gauge fields (in a manner we will discuss in Section 6), the axial transformation remains a symmetry of the classical Lagrangian. But it doesn't survive the quantization process. It is the archetypal example of an *anomaly*: a symmetry of the classical theory that is not preserved in the quantum theory.

4.7 Plane Wave Solutions

Let's now study the solutions to the Dirac equation

$$(i\gamma^\mu\partial_\mu - m)\psi = 0 \tag{4.103}$$

We start by making a simple ansatz:

$$\psi = u(\vec{p})\, e^{-ip\cdot x} \tag{4.104}$$

where $u(\vec{p})$ is a four-component spinor, independent of spacetime x which, as the notation suggests, can depend on the 3-momentum $\vec{p}$. The Dirac equation then becomes

$$(\gamma^\mu p_\mu - m)u(\vec{p}) = \begin{pmatrix} -m & p_\mu\sigma^\mu \\ p_\mu\bar{\sigma}^\mu & -m \end{pmatrix} u(\vec{p}) = 0 \tag{4.105}$$

where we're again using the definition,

$$\sigma^\mu = (1, \sigma^i) \quad \text{and} \quad \bar{\sigma}^\mu = (1, -\sigma^i) \tag{4.106}$$

Claim: The solution to (4.105) is

$$u(\vec{p}) = \begin{pmatrix} \sqrt{p\cdot\sigma}\,\xi \\ \sqrt{p\cdot\bar{\sigma}}\,\xi \end{pmatrix} \tag{4.107}$$

for any 2-component spinor ξ which we will normalize to $\xi^\dagger \xi = 1$.

Proof: Let's write $u(\vec{p})^T = (u_1, u_2)$. Then equation (4.105) reads

$$(p \cdot \sigma)\, u_2 = m u_1 \quad \text{and} \quad (p \cdot \bar{\sigma}) u_1 = m u_2 \tag{4.108}$$

Either one of these equations implies the other, a fact which follows from the identity $(p \cdot \sigma)(p \cdot \bar{\sigma}) = p_0^2 - p_i p_j \sigma^i \sigma^j = p_0^2 - p_i p_j \delta^{ij} = p_\mu p^\mu = m^2$. To start with, let's try the ansatz $u_1 = (p \cdot \sigma)\xi'$ for some spinor ξ'. Then the second equation in (4.108) immediately tells us that $u_2 = m\xi'$. So we learn that any spinor of the form

$$u(\vec{p}) = A \begin{pmatrix} (p \cdot \sigma)\, \xi' \\ m \xi' \end{pmatrix} \tag{4.109}$$

with constant A is a solution to (4.105). To make this more symmetric, we choose $A = 1/m$ and $\xi' = \sqrt{p \cdot \bar{\sigma}}\, \xi$ with constant ξ. Then $u_1 = (p \cdot \sigma)\sqrt{p \cdot \bar{\sigma}}\, \xi = m\sqrt{p \cdot \sigma}\, \xi$. So we get the promised result (4.107) $\square$

Negative Frequency Solutions

We get further solutions to the Dirac equation from the ansatz

$$\psi = v(\vec{p})\, e^{+ip\cdot x} \tag{4.110}$$

Solutions of the form (4.104), which oscillate in time as $\psi \sim e^{-iEt}$, are called positive frequency solutions. Those of the form (4.110), which oscillate as $\psi \sim e^{+iEt}$, are negative frequency solutions. It's important to note however that both are solutions to the classical field equations and both have positive energy (4.92). The Dirac equation requires that the 4-component spinor $v(\vec{p})$ satisfies

$$(\gamma^\mu p_\mu + m) v(\vec{p}) = \begin{pmatrix} m & p_\mu \sigma^\mu \\ p_\mu \bar{\sigma}^\mu & m \end{pmatrix} v(\vec{p}) = 0 \tag{4.111}$$

which is solved by

$$v(\vec{p}) = \begin{pmatrix} \sqrt{p \cdot \sigma}\, \eta \\ -\sqrt{p \cdot \bar{\sigma}}\, \eta \end{pmatrix} \tag{4.112}$$

for some 2-component spinor η which we take to be constant and normalized to $\eta^\dagger \eta = 1$.

4.7.1 Some Examples

Consider the positive frequency solution with mass m and 3-momentum $\vec{p} = 0$,

$$u(\vec{p}) = \sqrt{m} \begin{pmatrix} \xi \\ \xi \end{pmatrix} \tag{4.113}$$

where ξ is any 2-component spinor. Spatial rotations of the field act on ξ by (4.26),

$$\xi \to e^{+i\vec{\varphi}\cdot\vec{\sigma}/2}\xi \tag{4.114}$$

The 2-component spinor ξ defines the *spin* of the field. This should be familiar from quantum mechanics. A field with spin up (down) along a given direction is described by the eigenvector of the corresponding Pauli matrix with eigenvalue +1 (-1 respectively). For example, $\xi^T = (1, 0)$ describes a field with spin up along the z-axis. After quantization, this will become the spin of the associated particle. In the rest of this section, we'll indulge in an abuse of terminology and refer to the classical solutions to the Dirac equations as "particles", even though they have no such interpretation before quantization.

Consider now boosting the particle with spin $\xi^T = (1, 0)$ along the x^3 direction, with $p^\mu = (E, 0, 0, p)$. The solution to the Dirac equation becomes

$$u(\vec{p}) = \begin{pmatrix} \sqrt{p\cdot\sigma}\begin{pmatrix} 1 \\ 0 \end{pmatrix} \\ \sqrt{p\cdot\bar{\sigma}}\begin{pmatrix} 1 \\ 0 \end{pmatrix} \end{pmatrix} = \begin{pmatrix} \sqrt{E-p^3}\begin{pmatrix} 1 \\ 0 \end{pmatrix} \\ \sqrt{E+p^3}\begin{pmatrix} 1 \\ 0 \end{pmatrix} \end{pmatrix} \tag{4.115}$$

In fact, this expression also makes sense for a massless field, for which $E = p^3$. (We picked the normalization (4.107) for the solutions so that this would be the case). For a massless particle we have

$$u(\vec{p}) = \sqrt{2E}\begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} \tag{4.116}$$

Similarly, for a boosted solution of the spin down $\xi^T = (0, 1)$ field, we have

$$u(\vec{p}) = \begin{pmatrix} \sqrt{p\cdot\sigma}\begin{pmatrix} 0 \\ 1 \end{pmatrix} \\ \sqrt{p\cdot\bar{\sigma}}\begin{pmatrix} 0 \\ 1 \end{pmatrix} \end{pmatrix} = \begin{pmatrix} \sqrt{E+p^3}\begin{pmatrix} 0 \\ 1 \end{pmatrix} \\ \sqrt{E-p^3}\begin{pmatrix} 0 \\ 1 \end{pmatrix} \end{pmatrix} \xrightarrow{m\to 0} \sqrt{2E}\begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} \tag{4.117}$$

Helicity

The helicity operator is the projection of the angular momentum along the direction of momentum,

$$h = \tfrac{i}{2}\epsilon_{ijk}\hat{p}^i\, S^{jk} = \tfrac{1}{2}\hat{p}_i \begin{pmatrix} \sigma^i & 0 \\ 0 & \sigma^i \end{pmatrix} \tag{4.118}$$

where S^{ij} is the rotation generator given in (4.25). The massless field with spin $\xi^T = (1,0)$ in (4.116) has helicity $h = 1/2$: we say that it is *right-handed*. Meanwhile, the field (4.117) has helicity $h = -1/2$: it is *left-handed*.

4.7.2 Some Useful Formulae: Inner and Outer Products

There are a number of identities that will be very useful in the following section, regarding the inner (and outer) products of the spinors $u(\vec{p})$ and $v(\vec{p})$. It's firstly convenient to introduce a basis ξ^s and η^s, $s = 1,2$ for the two-component spinors such that

$$\xi^{r\,\dagger}\xi^s = \delta^{rs} \quad \text{and} \quad \eta^{r\,\dagger}\eta^s = \delta^{rs} \tag{4.119}$$

for example,

$$\xi^1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad \text{and} \quad \xi^2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \tag{4.120}$$

and similarly for η^s. Let's deal first with the positive frequency plane waves. The two independent solutions are now written as

$$u^s(\vec{p}) = \begin{pmatrix} \sqrt{p\cdot\sigma}\,\xi^s \\ \sqrt{p\cdot\bar{\sigma}}\,\xi^s \end{pmatrix} \tag{4.121}$$

We can take the inner product of four-component spinors in two different ways: either as $u^\dagger\cdot u$, or as $\bar{u}\cdot u$. Of course, only the latter will be Lorentz invariant, but it turns out that the former is needed when we come to quantize the theory. Here we state both:

$$\begin{aligned} u^{r\,\dagger}(\vec{p})\cdot u^s(\vec{p}) &= \left(\xi^{r\,\dagger}\sqrt{p\cdot\sigma}\,,\, \xi^{r\,\dagger}\sqrt{p\cdot\bar{\sigma}}\right)\begin{pmatrix} \sqrt{p\cdot\sigma}\,\xi^s \\ \sqrt{p\cdot\bar{\sigma}}\,\xi^s \end{pmatrix} \\ &= \xi^{r\,\dagger}p\cdot\sigma\xi^s + \xi^{r\,\dagger}p\cdot\bar{\sigma}\xi^s = 2\xi^{r\,\dagger}p_0\xi^s = 2p_o\delta^{rs} \end{aligned} \tag{4.122}$$

while the Lorentz invariant inner product is

$$\bar{u}^r(\vec{p})\cdot u^s(\vec{p}) = \left(\xi^{r\,\dagger}\sqrt{p\cdot\sigma}\,,\, \xi^{r\,\dagger}\sqrt{p\cdot\bar{\sigma}}\right)\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} \sqrt{p\cdot\sigma}\,\xi^s \\ \sqrt{p\cdot\bar{\sigma}}\,\xi^s \end{pmatrix} = 2m\delta^{rs} \tag{4.123}$$

We have analogous results for the negative frequency solutions, which we may write as

$$v^s(\vec{p}) = \begin{pmatrix} \sqrt{p\cdot\sigma}\,\eta^s \\ -\sqrt{p\cdot\bar{\sigma}}\,\eta^s \end{pmatrix} \quad \begin{array}{ll} \text{with} & v^{r\,\dagger}(\vec{p})\cdot v^s(\vec{p}) = 2p_0\delta^{rs} \\ \text{and} & \bar{v}^r(\vec{p})\cdot v^s(\vec{p}) = -2m\delta^{rs} \end{array} \tag{4.124}$$

We can also compute the inner product between u and v. We have

$$\begin{aligned} \bar{u}^r(\vec{p})\cdot v^s(\vec{p}) &= \left(\xi^{r\,\dagger}\sqrt{p\cdot\sigma}\,,\,\xi^{r\,\dagger}\sqrt{p\cdot\bar{\sigma}}\right)\gamma^0 \begin{pmatrix} \sqrt{p\cdot\sigma}\,\eta^s \\ -\sqrt{p\cdot\bar{\sigma}}\,\eta^s \end{pmatrix} \\ &= \xi^{r\dagger}\sqrt{(p\cdot\bar{\sigma})(p\cdot\sigma)}\eta^s - \xi^{r\,\dagger}\sqrt{(p\cdot\bar{\sigma})(p\cdot\sigma)}\eta^s = 0 \end{aligned} \tag{4.125}$$

and similarly, $\bar{v}^r(\vec{p})\cdot u^s(\vec{p}) = 0$. However, when we come to $u^\dagger\cdot v$, it is a slightly different combination that has nice properties (and this same combination appears when we quantize the theory). We look at $u^{r\,\dagger}(\vec{p})\cdot v^s(-\vec{p})$, with the 3-momentum in the spinor v taking the opposite sign. Defining the 4-momentum $(p')^\mu = (p^0,-\vec{p})$, we have

$$\begin{aligned} u^{r\,\dagger}(\vec{p})\cdot v^s(-\vec{p}) &= \left(\xi^{r\,\dagger}\sqrt{p\cdot\sigma}\,,\,\xi^{r\,\dagger}\sqrt{p\cdot\bar{\sigma}}\right) \begin{pmatrix} \sqrt{p'\cdot\sigma}\,\eta^s \\ -\sqrt{p'\cdot\bar{\sigma}}\,\eta^s \end{pmatrix} \\ &= \xi^{r\dagger}\sqrt{(p\cdot\bar{\sigma})(p'\cdot\sigma)}\eta^s - \xi^{r\,\dagger}\sqrt{(p\cdot\bar{\sigma})(p'\cdot\sigma)}\eta^s \end{aligned} \tag{4.126}$$

Now the terms under the square-root are given by $(p\cdot\bar{\sigma})(p'\cdot\sigma) = (p_0+p_i\sigma^i)(p_0-p_i\sigma^i) = p_o^2-\vec{p}^{\,2} = m^2$. The same expression holds for $(p'\cdot\bar{\sigma})(p\cdot\sigma)$, and the two terms cancel. We learn

$$u^{r\,\dagger}(\vec{p})\cdot v^s(-\vec{p}) = v^{r\,\dagger}(\vec{p})\cdot u^s(-\vec{p}) = 0 \tag{4.127}$$

Outer Products

There's one last spinor identity that we need before we turn to the quantum theory. It is:

Claim:

$$\sum_{s=1}^{2} u^s(\vec{p})\,\bar{u}^s(\vec{p}) = \not{p} + m \tag{4.128}$$

where the two spinors are not now contracted, but instead placed back to back to give a 4×4 matrix. Also,

$$\sum_{s=1}^{2} v^s(\vec{p})\,\bar{v}^s(\vec{p}) = \not{p} - m \tag{4.129}$$

Proof:

$$\sum_{s=1}^{2} u^s(\vec{p})\,\bar{u}^s(\vec{p}) = \sum_{s=1}^{2} \begin{pmatrix} \sqrt{p\cdot\sigma}\,\xi^s \\ \sqrt{p\cdot\bar{\sigma}}\,\xi^s \end{pmatrix} \left(\xi^{s\,\dagger}\sqrt{p\cdot\bar{\sigma}}\,,\, \xi^{s\,\dagger}\sqrt{p\cdot\sigma}\right) \tag{4.130}$$

But $\sum_s \xi^s\,\xi^{s\,\dagger} = \mathbf{1}$, the 2×2 unit matrix, which then gives us

$$\sum_{s=1}^{2} u^s(\vec{p})\bar{u}^s(\vec{p}) = \begin{pmatrix} m & p\cdot\sigma \\ p\cdot\bar{\sigma} & m \end{pmatrix} \tag{4.131}$$

which is the desired result. A similar proof works for $\sum_s v^s(\vec{p})\bar{v}^s(\vec{p})$. $\square$